Predicting effects of additive noise and hearing-instrument signal processing on consonant recognition and confusions

Zaar, Johannes; Dau, Torsten

Publication date:
2017

Document Version
Publisher's PDF, also known as Version of record

Predicting effects of additive noise and hearing-instrument signal processing on consonant recognition and confusions

Johannes Zaar[a)] and Torsten Dau

Hearing Systems Group, Department of Electrical Engineering, Technical University of Denmark, DK-2800 Kgs. Lyngby, Denmark

BACKGROUND AND OBJECTIVE

- Consonant-in-noise perception in normal-hearing (NH) listeners critically depends on fine details in the stimuli [Zaar and Dau, 2015].
- Consonant perception tests were shown to also be sensitive to effects of high-frequency amplification and non-linear frequency compression (NLFC) in hearing-impaired (HI) listeners [Schmitt *et al.*, 2016].
- Approaches to predict consonant-in-noise perception data [Cooke, 2006; Jürgens and Brand, 2009] showed:
  - Reasonable predictions of consonant recognition
  - Inaccurate predictions of consonant confusions
- Here, a consonant perception model that accounts for consonant recognition and confusions is proposed and evaluated in conditions of:
  - Stationary noise [Zaar and Dau, 2015]
  - Hearing-aid signal processing
  - Simulated cochlear-implant processing [DiNino *et al.*, 2016]

THE MODEL

- Extension of Dau *et al.* (1997) auditory perception model towards predicting microscopic speech perception
- A-priori knowledge about the correct speech token used in the template-matching procedure
- Variance σ^2_{int} of internal noise adjusted based on grand average recognition scores (model calibration)

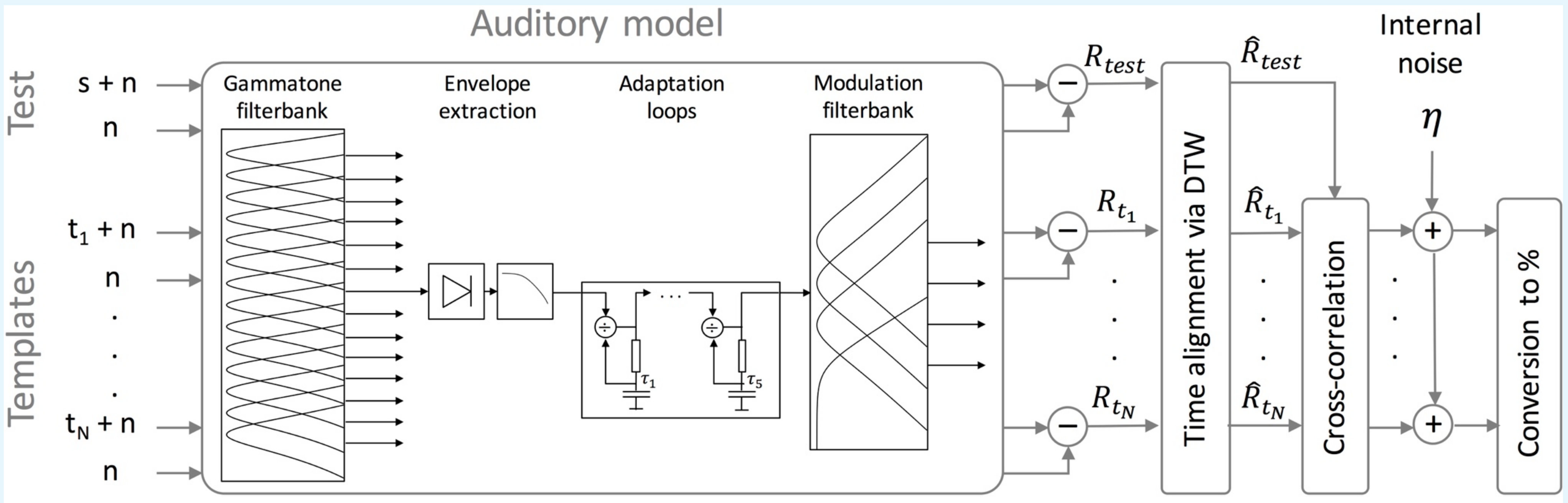

Fig. 1: Scheme of the proposed consonant perception model.

PREDICTING EFFECTS OF ADDITIVE STATIONARY NOISE

- Stimuli and data taken from Zaar and Dau (2015):
  - 15 consonant-vowels (CVs) in white noise
  - /bi, di, fi, gi, hi, ji, ki, li, mi, ni, pi, si, ʃi, ti, vi/
  - 6 speech tokens of each CV (3 by female / 3 by male speaker, both native speakers of Danish)
  - SNRs of -15, -12, -6, 0, 6, and 12 dB
  - 8 NH listeners (native speakers of Danish)
- Reference data for predictions: across-listener average data
- Internal-noise variance σ^2_{int} set to 0.05

Average consonant recognition

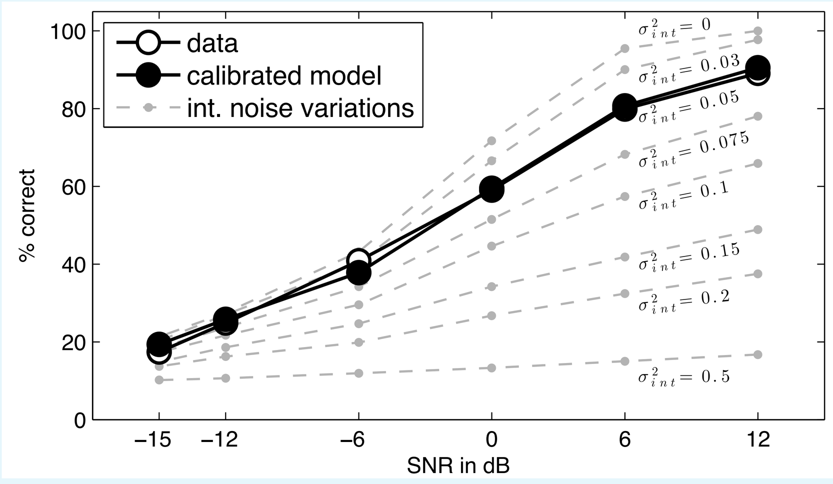

Fig. 2: Measured and predicted grand average consonant recognition scores

Detailed consonant recognition

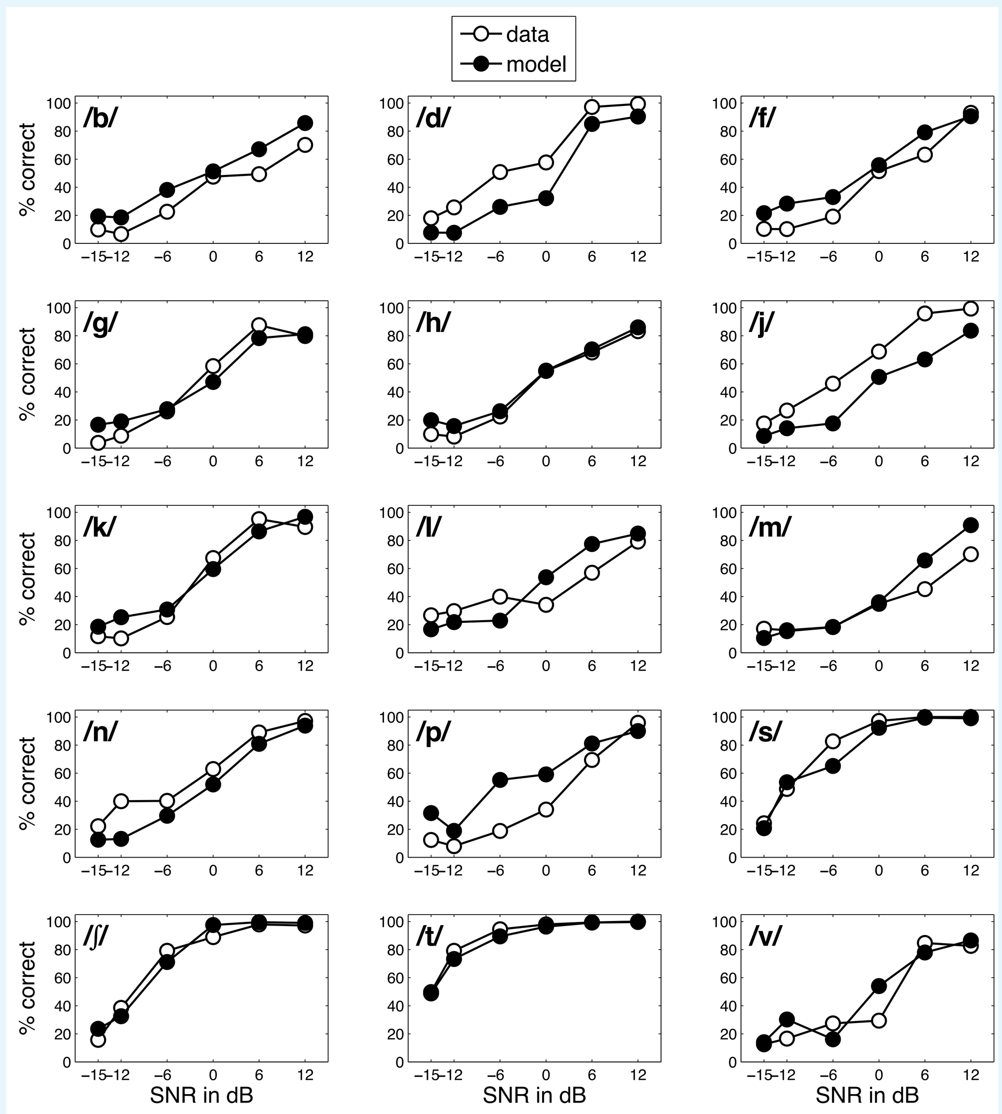

Fig. 3: Measured and predicted consonant-specific recognition scores

Consonant confusions

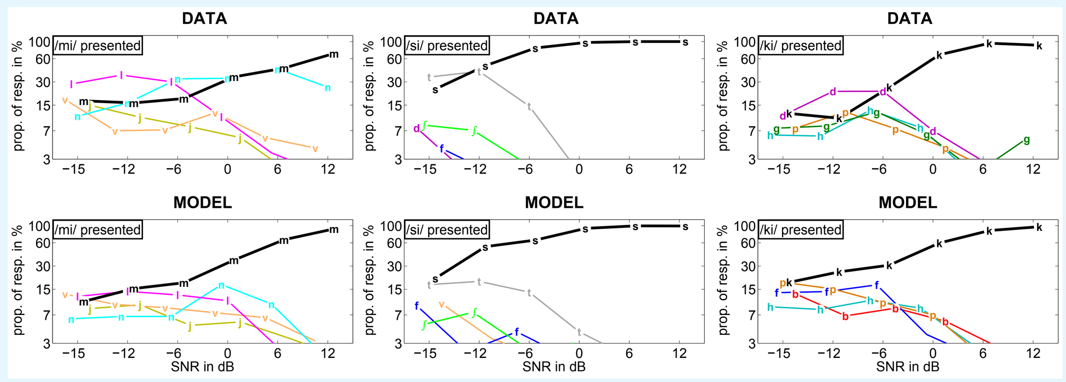

Fig. 4: Measured and predicted confusion patterns

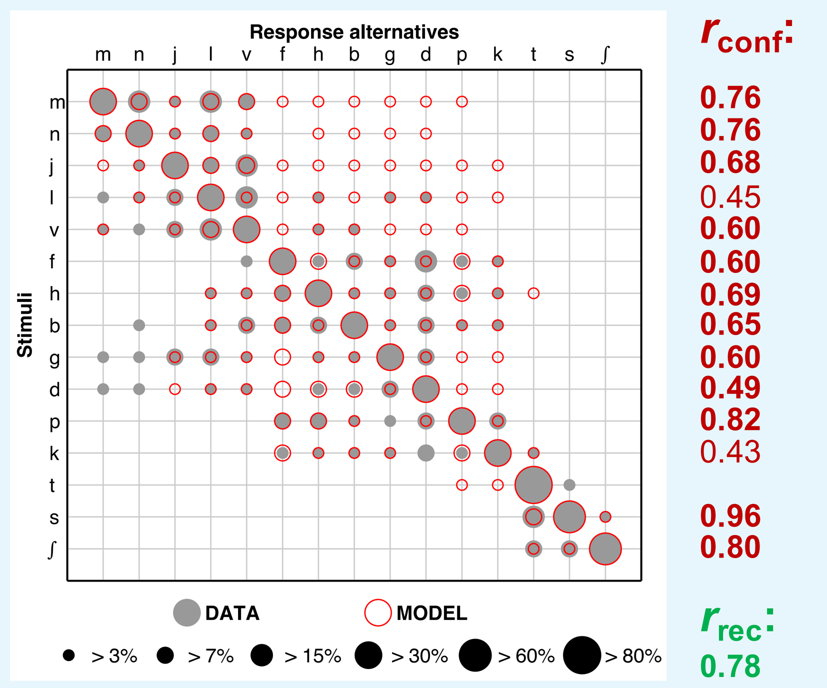

Fig. 5: Measured and predicted confusion matrix (averaged across SNR)

PREDICTING EFFECTS OF HEARING-AID SIGNAL PROCESSING

- 12 CVs /ba, ga, da, pa, ka, ta, sa6, sa9, ʃa3, ʃa5, fa, tsa/ (female speaker), mixed with speech-shaped noise at 8 dB SNR, presented to 10 NH listeners.
- Conditions obtained using Phonak Naída HAs on KEMAR: (i) ***unaided***, (ii) ***default***, (iii) non-linear frequency compression (***NLFC***), (iv) impulse noise suppression (***INS***), (v) ***NLFC&INS***.
- Recognition scores:

Unaided	Default	NLFC	INS	NLFC&INS
95.9%	93.7%	55.3%	92.3%	56.2%

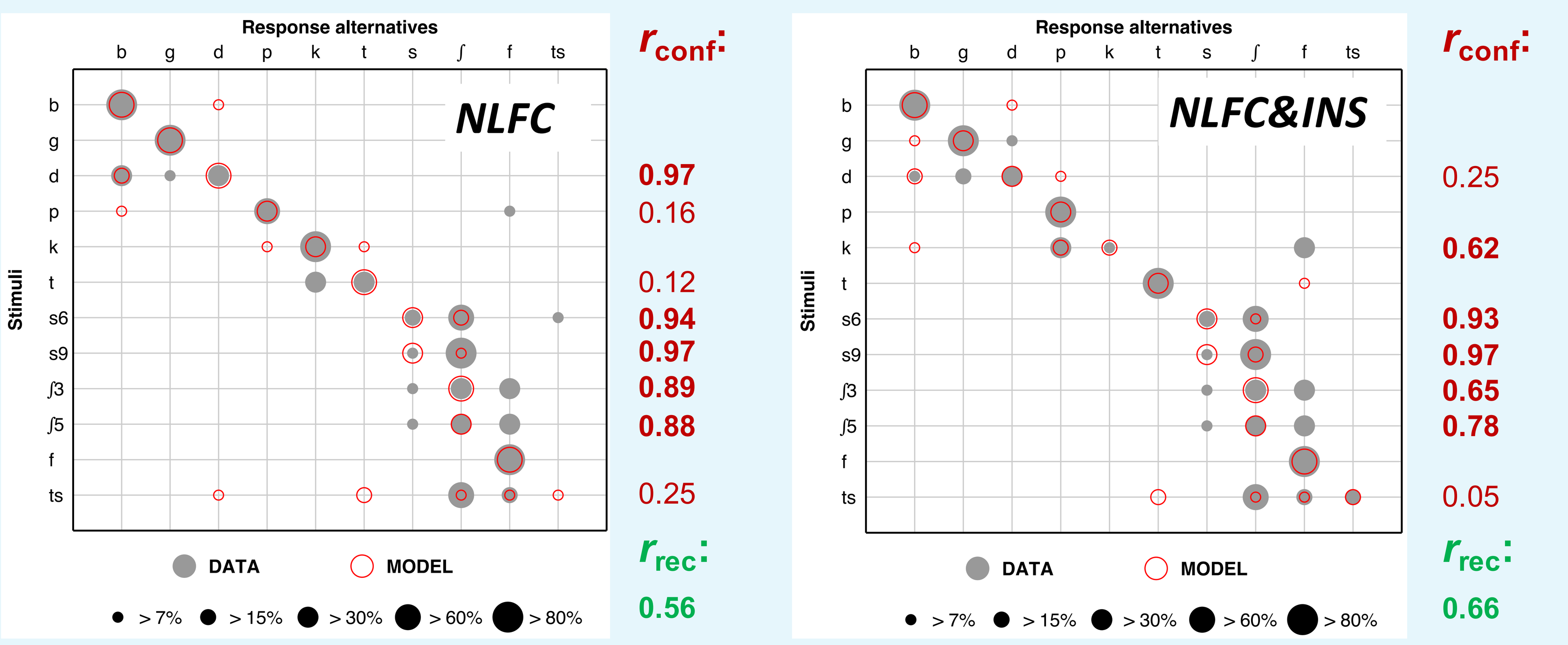

Fig. 6: Measured and predicted confusion matrices in the *NLFC* (left) and *NLFC&INS* (right) conditions.

PREDICTING EFFECTS OF COCHLEAR-IMPLANT SIGNAL PROCESSING

- Data from DiNino *et al.* (2016): 16 vowel-consonant-vowels (VCVs) /aba, aga, ada, apa, aka, ata, afa, ava, atha, asa, aza, asha, aja, ama, ana, ala/ in quiet, presented to 12 NH listeners.
- Cochlear-implant (CI) simulations obtained using appropriate noise-vocoding [Litvak *et al.*, 2007], see Fig. 7.

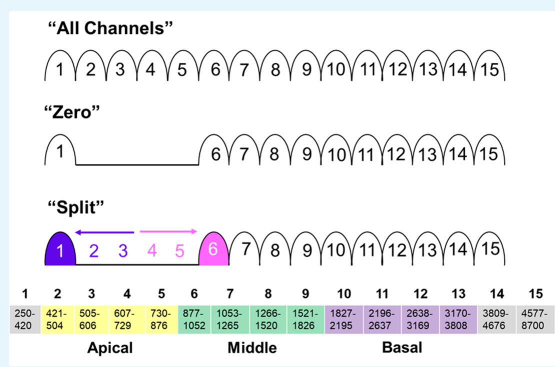

Fig. 7: CI simulation conditions

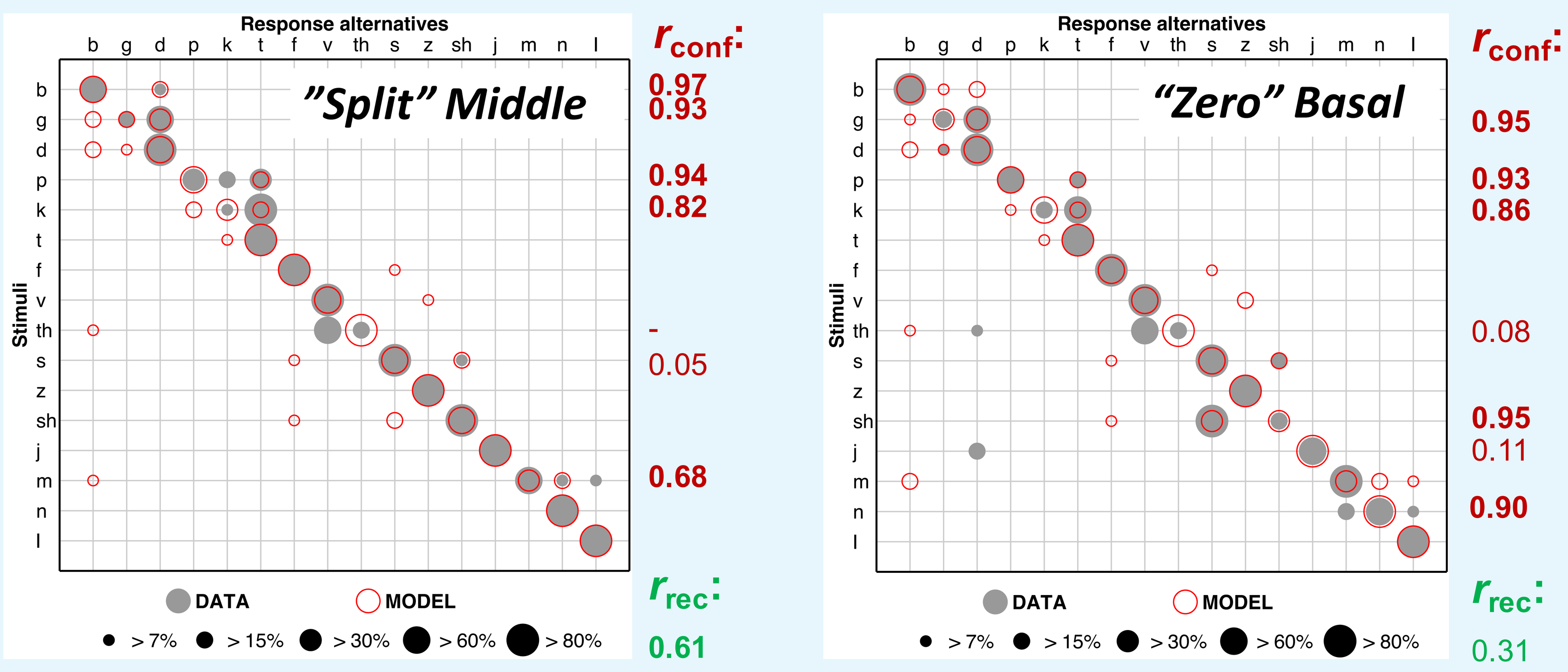

Fig. 8: Measured and predicted confusion matrices in the "Split" Middle (left) and *"Zero" Basal* (right) conditions.

CONCLUSIONS

- The proposed model accounts well for consonant recognition and confusions in all considered conditions, including effects induced by hearing-instrument processing.
- The model could therefore be useful for evaluating hearing-instrument processing strategies, particularly when combined with simulations of individual hearing impairment.

ACKNOWLEDGEMENTS

This research has been funded with support from the European Commission under Contract No. FP7-PEOPLE-2011-290000. We thank Ralph-Peter Derleth and Nicola Schmitt (Sonova AG) for their help with the HA signal processing experiment. We also thank Mishaela DiNino and Julie Bierer (University of Washington) for providing the stimuli and data from their CI study.

REFERENCES

Cooke, M. (2006): "A glimpsing model of speech perception in noise" , J. Acoust. Soc. Am., 119 (3): 1562–1573.
Dau, T., Kollmeier, B., and Kohlrausch, A. (1997): "Modeling auditory processing of amplitude modulation. I. Detection and Masking with narrow-band carriers" , J. Acoust. Soc. Am., 102 (5): 2892–2905.
DiNino, M., Wright, R. A., Winn, M. B., and Bierer, J. A. (2016): "Vowel and consonant confusions from spectrally manipulated stimuli designed to simulate poor cochlear implant electrode-neuron interfaces", J. Acoust. Soc. Am., 140 (6): 4404–4418.
Jürgens, T., and Brand, T. (2009): "Microscopic prediction of speech recognition for listeners with normal hearing in noise using an auditory model" J. Acoust. Soc. Am., 126 (5): 2635–2648.
Litvak, L. M., Spahr, A. J., Saoji, A. A., and Fridman, G. Y. (2007): "Relationship between perception of spectral ripple and speech recognition in cochlear implant and vocoder listeners" J. Acoust. Soc. Am., 122 (2): 982–991.
Schmitt, N., Winkler, A., Boretzki, M., and Holube, I. (2016): "A phoneme perception test method for high-frequency hearing aid fitting," J. Am. Acad. Audiol. 27 (5): 367–379.
Zaar, J., and Dau, T. (2015): "Sources of variability in consonant perception of normal-hearing listeners", J. Acoust. Soc. Am., 138 (3): 1253–1267.

a) jzaar@elektro.dtu.dk